Accurate Fault Prediction Using BlueGene/P System Logs via Geometric Reduction

Josh Thompson

Colorado State University

June 28, 2010

Introduction

From Raw Data to Fault Prediction

- We build two algorithms for fault prediction using raw system-log data.
- This work is prelimnary, and has only been applied to a limited dataset.
- However, the results seem promising.

Machine: Eugene.ccs.ornl.gov

2-rack BlueGene/P

Data obtained from directly from RAS system logs.

- Numeric Data
  - Seven Files Titled: Fan, Node, Lcard, Lcardp, Serv, Srvc, Bulk.
  - Each file represents a *component*.
- Text Data
  - Event Log: What happened, when and where.

Our Initial Dataset

- Comprises 25% of Eugene (512 nodes).
- Required 17 GB of Hard Drive space.

Machine: Eugene.ccs.ornl.gov

2-rack BlueGene/P

Data obtained from directly from RAS system logs.

- Numeric Data
  - Seven Files Titled: Fan, Node, Lcard, Lcardp, Serv, Srvc, Bulk.
  - Each file represents a *component*.
- Text Data
  - Event Log: What happened, when and where.

Our Initial Dataset

- Comprises 25% of Eugene (512 nodes).
- Required 17 GB of Hard Drive space.

Machine: Eugene.ccs.ornl.gov

2-rack BlueGene/P

Data obtained from directly from RAS system logs.

- Numeric Data
  - Seven Files Titled: Fan, Node, Lcard, Lcardp, Serv, Srvc, Bulk.
  - Each file represents a *component*.
- Text Data
  - Event Log: What happened, when and where.

Our Initial Dataset

- Comprises 25% of Eugene (512 nodes).
- Required 17 GB of Hard Drive space.

Machine: Eugene.ccs.ornl.gov

2-rack BlueGene/P

Data obtained from directly from RAS system logs.

- Numeric Data
  - Seven Files Titled: Fan, Node, Lcard, Lcardp, Serv, Srvc, Bulk.
  - Each file represents a *component*.
- Text Data
  - Event Log: What happened, when and where.

Our Initial Dataset

- Comprises 25% of Eugene (512 nodes).
- Required 17 GB of Hard Drive space.

BlueGene/P - RAS system

Raw Numeric data: Used for Training & Testing

Node Data

Timestamps	Location	MaxTemp1	MinTemp1	MaxTemp2	Volt12	Volt33	Volt50
1.4856e8	R00-M0-N5-J09	32	29	55	1.15	3.22	5.05
1.5356e8	R00-M0-N1-J01	0	29	56	1.17	3.32	5.07
1.5356e8	R00-M0-N7-J02	30	29	30	1.16	3.21	4.97
1.6546e8	R00-M0-N5-J02	32	29	14	1.13	3.20	5.03
1.6546e8	R00-M0-N5-J09	31	29	100	1.16	3.25	4.99
1.8454e8	R00-M0-N5-J03	30	29	55	1.16	3.32	5.06
1.4856e8	R00-M0-N5-J09	30	29	40	1.15	3.31	4.95

The Numerical Files Can Be Very Dirty

BlueGene/P - RAS system

Raw Numeric data: Used for Training & Testing

Node Data

Timestamps	Location	MaxTemp1	MinTemp1	MaxTemp2	Volt12	Volt33	Volt50
1.4856e8	R00-M0-N5-J09	32	0	55	1.15	3.22	5.05
1.5356e8	R00-M0-N1-J01	30	0	56	1.17	3.32	5.07
1.5356e8	R00-M0-N7-J02	30	0	30	1.16	3.21	4.97
1.6546e8	R00-M0-N5-J02	32	0	14	1.13	3.20	5.03
1.6546e8	R00-M0-N5-J09	31	0	20	1.16	3.25	4.99
1.8454e8	R00-M0-N5-J03	30	0	55	1.16	3.32	5.06
1.4856e8	R00-M0-N5-J09	30	0	40	1.15	3.31	4.95

Multiple sub-components report at each Timestamp.

BlueGene/P - RAS system

Raw Numeric data: Used for Training & Testing

Node Data

Timestamps	Location	MaxTemp1	MinTemp1	MaxTemp2	Volt12	Volt33	Volt50
1.4856e8	R00-M0-N5-J09	32	29	55	1.15	3.22	5.05
1.5356e8	R00-M0-N1-J01	30	27	56	1.17	3.32	5.07
1.5356e8	R00-M0-N7-J02	30	26	30	1.16	3.21	4.97
1.6546e8	R00-M0-N5-J02	32	23	14	1.13	3.20	5.03
1.6546e8	R00-M0-N5-J09	31	29	20	1.16	3.25	4.99
1.8454e8	R00-M0-N5-J03	30	26	55	1.16	3.32	5.06
1.4856e8	R00-M0-N5-J08	30	25	40	1.15	3.31	4.95

Sample Rates: Roughly once every 5 - 10 minutes.

BlueGene/P - RAS system

Raw Textual data: Used for Ground-Truth

Event-Log

Timestamps	Location	Severity	Component
1.5856e8	ROO-B-P2	WARN	MMCS
1.7356e8	R00-M0-A9	ERROR	BARMETAL
1.7356e8	R00-M0-N1-J06	FAULT	KERNEL
1.8546e8	ROO-B-P3	ERROR	KERNEL
1.8346e8	R00-M0-N1-J05	FAULT	MMCS
1.8454e8	R00-M0-L0-U01	UNKNOWN	CARD
1.8589e8	R00-M0-A1	WARN	KERNEL

We predict occurances of FAULT in the Event-Log.

BlueGene/P - RAS system

Raw Textual data: Used for Ground-Truth

Event-Log

Timestamps	Location	Severity	Component
1.5856e8	ROO-B-P2	WARN	MMCS
1.7356e8	R00-M0-A9	ERROR	BARMETAL
1.7356e8	R00-M0-N1-J06	FAULT	KERNEL
1.8546e8	ROO-B-P3	ERROR	KERNEL
1.8346e8	R00-M0-N1-J05	FAULT	MMCS
1.8454e8	R00-M0-L0-U01	UNKNOWN	CARD
1.8589e8	R00-M0-A1	WARN	KERNEL

We predict occurances of FAULT in the Event-Log.

BlueGene/P - RAS system

Raw Textual data: Used for Ground-Truth

Event-Log

Timestamps	Location	Severity	Component
1.5856e8	ROO-B-P2	WARN	MMCS
1.7356e8	R00-M0-A9	ERROR	BARMETAL
1.7356e8	R00-M0-L0-U01	FAULT	KERNEL
1.8546e8	ROO-B-P3	ERROR	KERNEL
1.8346e8	R00-M0-A8	FAULT	MMCS
1.8454e8	R00-M0-N1-J06	UNKNOWN	CARD
1.8589e8	R00-M0-N1-J06	WARN	KERNEL

Twelve faults occured in six tight clusters (.2 secs - 40 minutes).

BlueGene/P - RAS system

Raw Textual data: Used for Ground-Truth

Event-Log

Timestamps	Location	Severity	Component
1.5856e8	ROO-B-P2	WARN	MMCS
1.7356e8	R00-M0-A9	ERROR	BARMETAL
1.7356e8	R00-M0-L0-U01	FAULT	KERNEL
1.8546e8	ROO-B-P3	ERROR	KERNEL
1.8346e8	R00-M0-A8	FAULT	MMCS
1.8454e8	R00-M0-N1-J06	UNKNOWN	CARD
1.8589e8	R00-M0-N1-J06	WARN	KERNEL

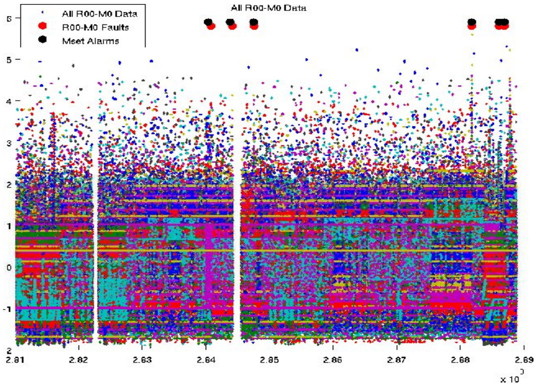
Twelve faults occured in six tight clusters (.2 secs - 40 minutes).

Data Cleaning & Reduction

All Normalized Data for R00-M0

Seven numeric files give seven matrices with varying rows and columns.

Interpolation is performed to weave the timeseries together.

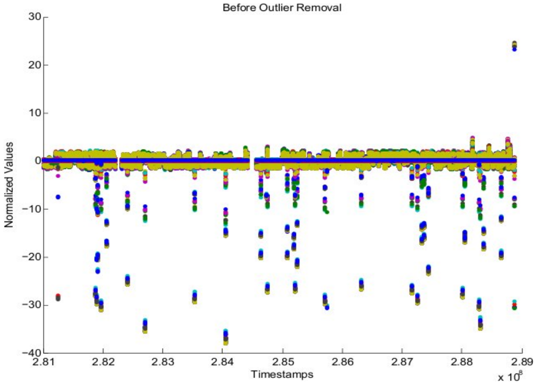


Data Cleaning & Reduction

Reduce Data to Bulk & Node

Obvious outliers are removed.

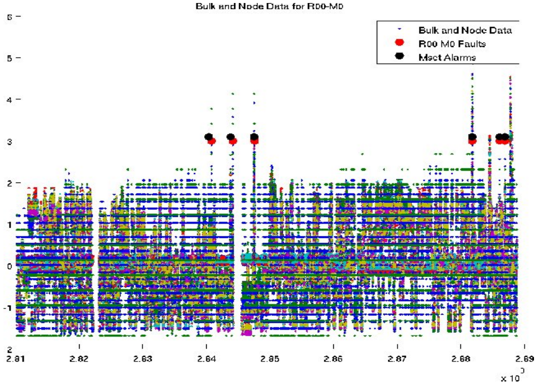
Here data falling 4 standard deviations away from mean are removed.



Data Cleaning & Reduction
Reduce Data to Bulk & Node

The correlation of data to faults is visible.

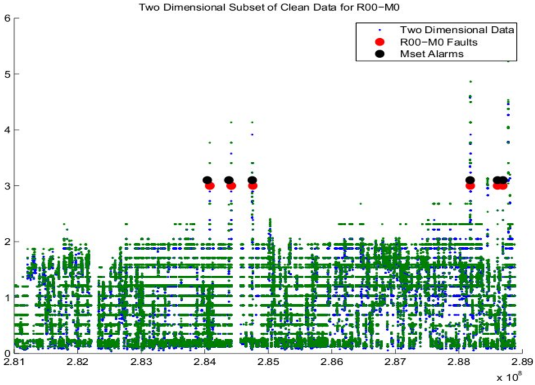
However further analysis is needed.



Data Cleaning & Reduction

Reduce Data to Bulk & Node

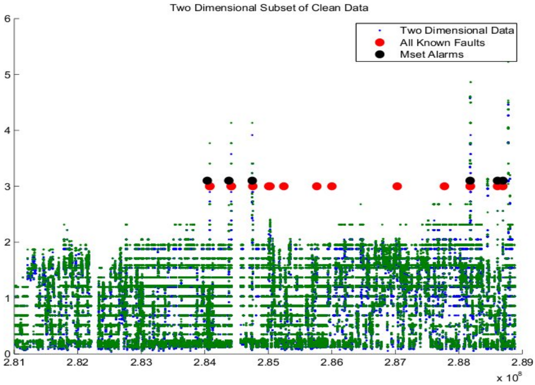
Data is mapped to $\mathbb{R}^2$ by choosing the top two extreme values at each timestamp



Data Cleaning & Reduction

Reduce Data to Bulk & Node

Little, but some, bleeding from adjacent equipment is observed



Our Fault Prediction Algorithms

Exploiting the Geometry of Data

- Two Distinct Approaches
  - MSET - Multivariate State Estimation Technique
  - Non-Negative Matrix Factorization (GSVD, and GLDA)
- In Both Cases
  - Algorithms detect geometric changes in data before faults occur.
  - Low dimensional data is used for prediction.

Our Fault Prediction Algorithms

Exploiting the Geometry of Data

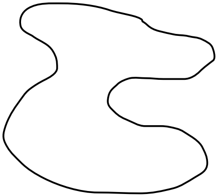
- Two Distinct Approaches
  - MSET - Multivariate State Estimation Technique
  - Non-Negative Matrix Factorization (GSVD, and GLDA)
- In Both Cases
  - Algorithms detect geometric changes in data before faults occur.
  - Low dimensional data is used for prediction.

Extending a Mapping of the Identity

Novelty detection

Assumption data lies on a special subset $S \subset X$ which has intrinsic structure: geometry, topology.

(Think: $X = \mathbb{R}^2$, and S is curve shown.)



We seek a mapping $\mathbf{f} : X \to X$ which preserves the intrinsic structure of S.

Extending a Mapping of the Identity

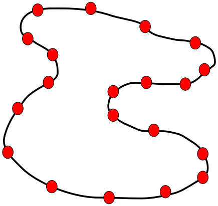
Novelty detection

Samples D from S are used to define a smooth mapping

$$\mathbf{f} : X \to X$$

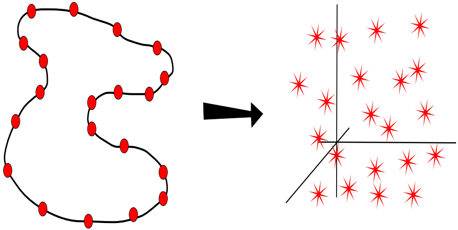
where $\mathbf{f}(D) = D$ and acts non-trivially elsewhere on X.

Smoothness guarantees that x near D is perturbed minimally.



Extending a Mapping of the Identity
Novelty detection

The mapping **f** factors through a similarity map Φ which scores the data, expressing self-similarity



Extending a Mapping of the Identity

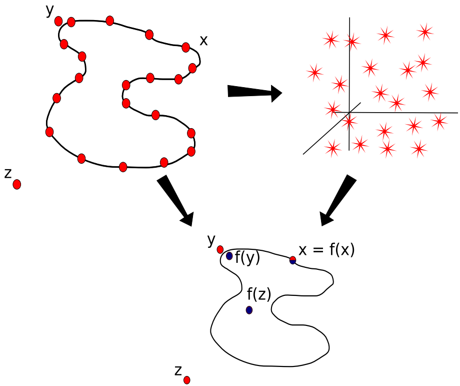
Novelty detection

The scores are used to reconstruct patterns

$\{f(x), f(y), f(z)\} \in X$

from the given patterns

$\{x, y, z\} \in X$.



Anamoly Detection via Function Evaluation

Build a Yes-No Function and Evaluate

- Encode entire library into one smooth mapping $f : \mathbb{R}^N \rightarrow \mathbb{R}^N$.
- If X is in L then $f(X) = X$.
- If X is NOT in L then $f(X) \neq X$.
- The defect between X and its reconstruction under f gives a measure of novelty.

Anamoly Detection via Function Evaluation

Can Detect Novelties It Has Not Seen Before

- A function evaluation can be performed in real-time.
- No signature required: It does not take one to know one.
- Previously unknown types of novelty can be detected.

Multi-Variate State Estimation Technique

- Memory matrix X of size (m, n). Contains data from m sensors, n samples each.
- The i^{th} column is an observation vector $X^{(i)}$ of the system at time i.
- Given a new pattern P construct a feature vector W expressing similarity between each $X^{(i)}$ and P.

$$W \equiv W(P) = (X^T \star X)^{-1}(X^T \star P).$$

- The matrix $X^T \star Y$ expresses the similarity of the given pattern Y with each sample in the memory.

The Kernel Mapping

Definition

Given $X \in R^m \times R^n$ and $Y \in R^m \times R^k$, we define $X \star Y \in R^n \times R^k$ as the matrix whose (i, j) coordintate is given by

$$X \star Y_{(i,j)} = 1 - \frac{\| X^{(i)} - Y^{(j)} \|^2}{\| X^{(i)} \|^2 + \| Y^{(j)} \|^2}.$$

The MSET mapping Φ_X defined as

$$\Phi_X(P) \equiv X * W(P)$$

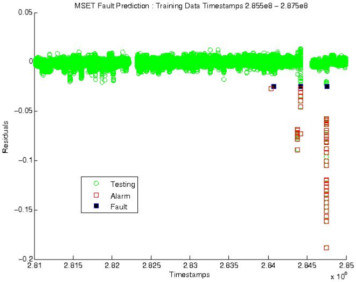
and the *Residual* is then

$$\mathcal{R} = \| \Phi_X(P) - P \| .$$

Fault Prediction

Using MSET residuals with a standard thresholding.

FAULT	LEADTIME
1	10.4 hours
2	10.2 hours
3	2.5 hours

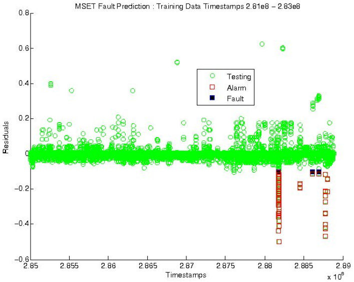


Here we use the latter data to predict earlier three faults.

Fault Prediction

Using MSET residuals with a standard thresholding.

FAULT	LEADTIME
4	1 hour
5	7 min.
6	-22 hours

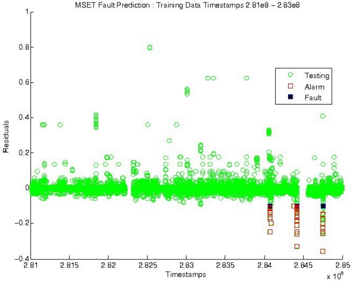


Here we use the earlier data to predict latter three faults.

Fault Prediction

Using MSET residuals with a standard thresholding.

FAULT	LEADTIME
1	-15 min.
2	10.2 hours
3	2.5 hours

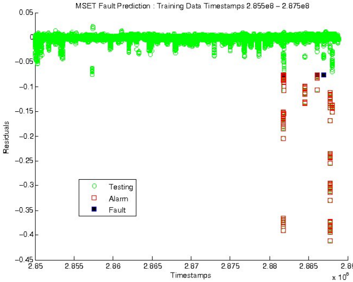


Here we use the earlier data to predict earlier three faults.

Fault Prediction

Using MSET residuals with a standard thresholding.

FAULT	LEADTIME
4	55 hour
5	13 min.
6	-22 hours

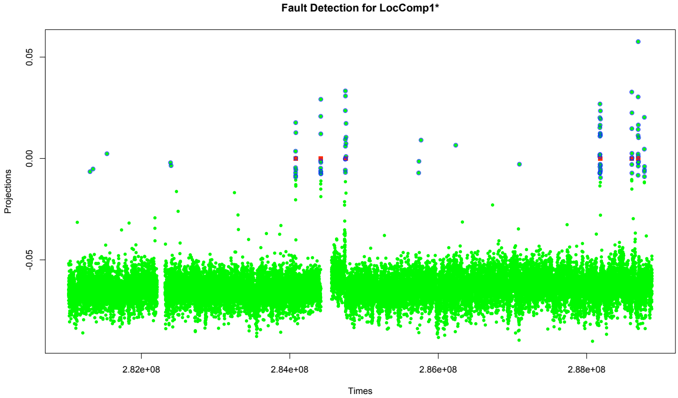


Here we use the latter data to predict latter three faults.

Fault Prediction

Using an affine transformation and windowed thresholding.

FAULT	LEADTIME
1	-2.5 min.
2	5 min.
3	80 min.

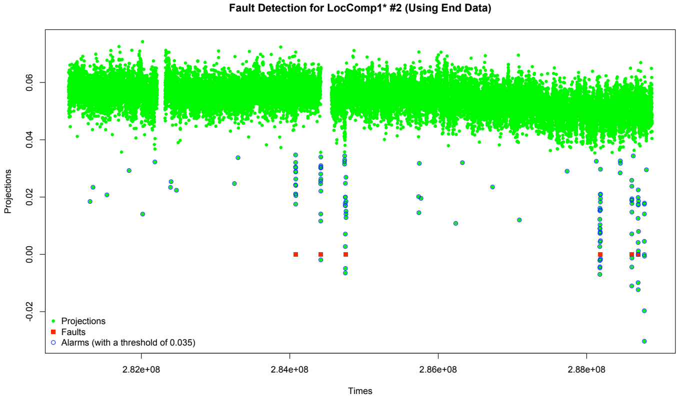


Here we use the earlier data to predict earlier three faults.

Fault Prediction

Using an affine transformation and windowed thresholding.

FAULT	LEADTIME
4	32.5 min.
5	10 min.
6	-2.5 hours



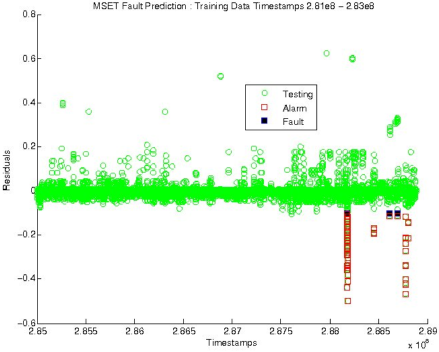
Here we use the earlier data to predict latter three faults.

Novelty Detection via MSET & NMF Residuals

ORNL supercomputer data

Our analysis had a false positive at February 24 2009 at 07:45.

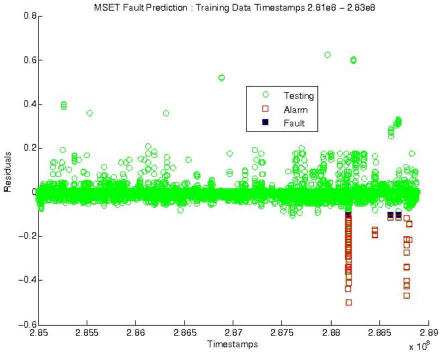
However, ...



Quote From BlueGeneP System Adminstrator

Our analysis outperformed a System Admin & caught a failure NOT NOTED in the logs

"...Not long after that (8:14 on 02/24/2009) I ran diagnostics on midplane R00-M0, and two nodes failed the tests and were put into service mode to be replaced."



Conclusions

From Raw Data to Fault Prediction

- Obtaining useful data from RAS-logs is challenging.
- Extracting concentrated information improves efficiency and accuracy.
- Function evaluation algorithms are fast and lend well to scaling.

Questions?

Thanks for your time!

Josh Thompsonon - Colorado State University
thompson@math.colostate.edu

Contributors:

Michael Kirby: CSU
Terry L. Jones: ORNL
Joshua Ladd: ORNL
David W. Dreisigmeyer: Univ. Pittsburgh

Special Thanks To:

Oak Ridge National Laboratory
Directorate of Central Intelligence Postdoctoral Program